DOCUMENT RESUME

ED 449 679 FL 026 581

AUTHOR Nakamura, Yuji
TITLE Rasch Based Analysis of Reading Ability Questionnaire.
PUB DATE 2000-03-00
NOTE 18p.; Research supported by Tokyo Keizai University, Research Grant APC 01-97. Published by International Christian University.
PUB TYPE Journal Articles (080) -- Reports - Research (143)
LANGUAGE English, Japanese
JOURNAL CIT Educational Studies; v42 p261-276 Mar 2000
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Students; *Communicative Competence (Languages); Comparative Analysis; *English (Second Language); Evaluation; *Factor Analysis; Foreign Countries; Higher Education; Language Teachers; Questionnaires; *Reading Comprehension; *Reading Tests; Second Language Instruction; Second Language Learning; Teacher Attitudes; Test Validity
IDENTIFIERS *Authentic Materials; Japan; Rasch Scaled Scores

ABSTRACT

This paper examines the results of a questionnaire on reading ability in English by Japanese college students, which was formerly analyzed using raw scores, from the viewpoint of Rasch measured scores. In the Rasch analysis, the basic requirements for measuring are the following: (1) reduction of experience to one dimensional abstraction; (2) comparisons among persons and items; (3) the linear magnitude inherent in positioning objects along a line; and (4) a unit determined by a process that can be repeated without modification over the range of the variable. With these requirements in mind, this article deals with the rating scale data by focusing on the idea of linearity. Results are examined along a continuum. It is concluded that reading for testing has two components: contents and skills. Contents should be authentic materials because students use skills to comprehend or draw conclusions from the contents. The eventual scores given by the tests are the end product of their comprehension. Thus the process of reading is analyzed by examining three factors: students are given authentic materials; they struggle to comprehend the material and try to use it; their scores are calculated. (6 references.) (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 449 679

教育研究
Educational Studies
42
March 2000

Rasch Based Analysis of Reading Ability Questionnaire

Yuji Nakamura

ラッシュモデルを用いた英語リーディング能力の分析

中村優治

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Prof. Yuji Nakamura
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE
2

国際基督教大学
International Christian University

抜刷 / offprint

International Christian University
教育研究 Educational Studies 42 / 2000
pp. 261-276
研究論文 Articles

Rasch Based Analysis of Reading Ability Questionnaire

ラッシュモデルを用いた英語リーディング能力の分析

Yuji Nakamura Tokyo Keizai University
中村優治 東京経済大学

Keywords READING ABILITY, TESTING, RASCH MODEL, FACTOR ANALYSIS
リーディング能力、テスティング、ラッシュモデル、因子分析

要旨

本稿はラッシュモデルのその統計的分析手法を利用することにより、英語のリーディング能力の構成概念をより科学的に、解明しようとするものである。使用したデータは中村 (1998) のアンケート調査結果によるもので、すでに素点による分析は終了している。今回の研究では、ラッシュモデルの linearity の考え方をもとに英語教育者のリーディング能力に対する潜在的な考え方と実際にデータに表されたものとの関係についてラッシュ得点を用いて分析を行った。結論として、contents と skills の 2 つの構成要素が確認され、英語のリーディングテストの作成に新たな示唆を与えたものと考えられる。また、古典的テスト理論との比較も行われており、さ

らなる研究が期待される。

This paper examines the results of the questionnaire on reading ability, which was formerly analyzed in Nakamura (1998) using raw scores, from the viewpoint of the Rasch measured scores. In the Rasch analysis, the basic requirements for measuring are as follows:

1) the reduction of experience to a one dimensional abstraction
2) more or less comparisons among persons and items
3) the idea of linear magnitude inherent in positioning objects along a line
4) a unit determined by a process which can be repeated without modification over the range of the variable (Wright and Masters 1982, p. 3)

In the present research, with these requirements in mind, we deal with the rating scale data by focusing on the idea of linearity. On the continuum of the scale, we look at the teachers' latent idea towards items (in the usual case "item difficulty"), and the teachers' expressed attitude towards items (in the usual case "person ability"). On the same continuum we look at the probability of 50% between the latent idea and the expressed attitude.

Nakamura (1998) was able to find six factors using raw scores, and concluded that we need to reconsider the relationship of each factor in practice depending on the situation of the individual institution.

Our present research attempts to provide a different result of the factor analysis, and attempts to question previous conclusions. In addition, the difference between Japanese and Non-Japanese teachers towards the items will be reported.

1. Factor analysis results

First, let us look at the factor analysis results. Tables 1-5 below show the results of factor analysis. The first factor in Table 1, which is composed (if we take items with factor loading over .30) of items 17, 15, 18, 16, 9 and 27, can be called Authentic Material Reading Ability. In particular, in the factor plot, A, B, C (17, 15, 18) cluster distinctively and contribute to this factor. All of the items included in this factor deal with contents of reading material, yet they should be presented as they are without being adjusted or modified for less able students. Item 27 (comprehension) was intentionally eliminated from this factor naming process, and is put into another factor.

Table 1

Factor 1 from principal component analysis of standardized residual correlations for items (sorted by loading)

Factor 1 explains 3.34 of 24 variance units

FACTOR	LOADING	MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER	ITEM
1	.69	59.4	1.08	1.07	A 17	journals
1	.60	60.3	1.04	1.03	B 15	editorials
1	.59	59.9	.97	.96	C 18	literature
1	.41	50.7	.84	.82	D 16	articles
1	.37	60.1	.80	.79	E 9	academic
1	.33	51.9	.78	.76	F 27	comprehensn
1	.25	58.7	.77	.77	G 2	predict
1	.21	51.6	1.08	1.10	H 10	authentic
1	.14	52.1	1.37	1.37	I 7	schemata
1	.11	27.8	1.01	1.09	J 19	main ideas
1	.05	43.3	1.07	1.01	K 25	topic sntnc
1	−.65	53.8	1.08	1.08	a 6	grammar
1	−.56	55.8	1.22	1.20	b 4	word
1	−.50	49.9	.94	.91	c 5	vocabulary
1	−.40	46.5	1.15	1.11	d 24	scanning
1	−.39	49.4	.96	.99	e 8	sentence
1	−.37	42.3	1.24	1.06	f 23	skimming
1	−.31	47.1	.99	1.05	g 3	discourse
1	−.16	43.3	.72	.70	h 26	conclusions
1	−.15	56.5	.83	.86	i 20	idioms
1	−.11	47.9	1.02	1.04	j 21	quick read
1	−.08	49.7	1.10	1.20	k 14	sum spokEng
1	−.02	48.9	.97	.93	l 13	sum writEng
1	.00	33.1	.98	.82	L 1	grasp

The second factor in Table 2, which consists of (items with factor loading over .30) items 24, 23, 10, 16, 2, 25 and 27, can be termed Skimming and Scanning Ability. The reason being that the first two items 24 and 23 make a great contribution to this factor as seen in the factor plot. Although it is correct to view these two items as skimming and scanning as skills for reading rather than as ability, it is also true that we can improve these skills by practice, and measure them as an ability on a scale. Furthermore, the item "authentic" is complicatedly involved in this factor because the skimming and scanning factor demands authentic materials.

Table 2

Factor 2 from principal component analysis of standardized residual correlations for items (sorted by loading)

Factor 2 explains 2.96 of 24 variance units

FACTOR	LOADING	MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER	ITEM
2	.63	46.5	1.15	1.11	d 24	scanning
2	.47	42.3	1.24	1.06	f 23	skimming
2	.41	51.6	1.08	1.10	H 10	authentic
2	.35	50.7	.84	.82	D 16	articles
2	.34	58.7	.77	.77	G 2	predict
2	.33	43.3	1.07	1.01	K 25	topic sntnc
2	.32	51.9	.78	.76	F 27	comprehensn
2	.29	27.8	1.01	1.09	J 19	main ideas
2	.28	52.1	1.37	1.37	I 7	schemata
2	.21	47.9	1.02	1.04	j 21	quick read
2	.20	33.1	.98	.82	L 1	grasp
2	.19	43.3	.72	.70	h 26	conclusions
2	.16	49.9	.94	.91	c 5	vocabulary
2	−.61	49.7	1.10	1.20	k 14	sum spokEng
2	−.46	48.9	.97	.93	l 13	sum writEng
2	−.42	59.9	.97	.96	C 18	literature
2	−.38	56.5	.83	.86	i 20	idioms
2	−.37	49.4	.96	.99	e 8	sentence
2	−.35	55.8	1.22	1.20	b 4	word
2	−.27	53.8	1.08	1.08	a 6	grammar
2	−.27	60.1	.80	.79	E 9	academic
2	−.23	60.3	1.04	1.03	B 15	editorials
2	−.21	59.4	1.08	1.07	A 17	journals
2	−.12	47.1	.99	1.05	g 3	discourse

The third factor in Table 3, which is composed of items 1, 13, 21, 26, 25 (items with factor loading over .30), can be named Prompt Context-Grasping Ability, because students must grasp the context of the reading material rather quickly, and draw some conclusions.

Table 3

Factor 3 from principal component analysis of standardized residual correlations for items (sorted by loading)

Factor 3 explains 2.44 of 24 variance units

FACTOR	LOADING	MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER	ITEM
3	.55	33.1	.98	.82	L 1	grasp
3	.47	48.9	.97	.93	l 13	sum writEng
3	.46	47.9	1.02	1.04	j 21	quick read
3	.41	43.3	.72	.70	h 26	conclusions
3	.35	43.3	1.07	1.01	K 25	topic sntnc
3	.26	58.7	.77	.77	G 2	predict
3	.25	47.1	.99	1.05	g 3	discourse
3	.22	60.3	1.04	1.03	B 15	editorials
3	.18	42.3	1.24	1.06	f 23	skimming
3	.13	49.7	1.10	1.20	k 14	sum spokEng
3	.11	56.5	.83	.86	i 20	idioms
3	.11	46.5	1.15	1.11	d 24	scanning
3	.07	59.4	1.08	1.07	A 17	journals
3	−.52	51.6	1.08	1.10	H 10	authentic
3	−.45	52.1	1.37	1.37	I 7	schemata
3	−.42	49.9	.94	.91	c 5	vocabulary
3	−.37	49.4	.96	.99	e 8	sentence
3	−.36	51.9	.78	.76	F 27	comprehensn
3	−.36	53.8	1.08	1.08	a 6	grammar
3	−.25	60.1	.80	.79	E 9	academic
3	−.17	50.7	.84	.82	D 16	articles
3	−.14	59.9	.97	.96	C 18	literature
3	−.06	55.8	1.22	1.20	b 4	word
3	−.05	27.8	1.01	1.09	J 19	main ideas

The fourth factor in Table 4, which consists of items 14, 20, 23, 19, 13, can be called Main Idea Summarizing Ability. This name is derived from the fact that students are required to grasp the main idea, and often asked to give a summary in spoken or written English. Item 23 can be eliminated from this factor since it makes more of a contribution to the second factor.

Table 4

Factor 4 from principal component analysis of standardized residual correlations for items (sorted by loading)

Factor 4 explains 2.07 of 24 variance units

FACTOR	LOADING	MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER	ITEM
4	.54	49.7	1.10	1.20	k 14	sum spokEng
4	.48	56.5	.83	.86	i 20	idioms
4	.44	42.3	1.24	1.06	f 23	skimming
4	.40	27.8	1.01	1.09	J 19	main ideas
4	.38	48.9	.97	.93	l 13	sum writEng
4	.25	51.6	1.08	1.10	H 10	authentic
4	.14	51.9	.78	.76	F 27	comprehensn
4	.09	47.9	1.02	1.04	j 21	quick read
4	.07	50.7	.84	.82	D 16	articles
4	.04	52.1	1.37	1.37	I 7	schemata
4	.02	59.9	.97	.96	C 18	literature
4	.01	46.5	1.15	1.11	d 24	scanning
4	−.51	58.7	.77	.77	G 2	predict
4	−.41	60.1	.80	.79	E 9	academic
4	−.37	33.1	.98	.82	L 1	grasp
4	−.35	55.8	1.22	1.20	b 4	word
4	−.31	47.1	.99	1.05	g 3	discourse
4	−.22	53.8	1.08	1.08	a 6	grammar
4	−.21	49.4	.96	.99	e 8	sentence
4	−.20	59.4	1.08	1.07	A 17	journals
4	−.12	43.3	.72	.70	h 26	conclusions
4	−.07	49.9	.94	.91	c 5	vocabulary
4	−.07	60.3	1.04	1.03	B 15	editorials
4	−.04	43.3	1.07	1.01	K 25	topic sntnc

The fifth factor is composed of items 27, 26, and 10, with the first two items having the greater influence. This factor can be termed Overall Comprehension Ability because the eventual goal of reading is to comprehend the given material, and draw some conclusions from the material.

Table 5

Factor 5 from principal component analysis of standardized residual correlations for items (sorted by loading)

Factor 5 explains 1.74 of 24 variance units

FACTOR	LOADING	MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER	ITEM
5	.56	51.9	.78	.76	F 27	comprehensn
5	.52	43.3	.72	.70	h 26	conclusions
5	.38	51.6	1.08	1.10	H 10	authentic
5	.24	43.3	1.07	1.01	K 25	topic sntnc
5	.23	55.8	1.22	1.20	b 4	word
5	.22	47.1	.99	1.05	g 3	discourse
5	.20	48.9	.97	.93	l 13	sum writEng
5	.17	33.1	.98	.82	L 1	grasp
5	.15	59.9	.97	.96	C 18	literature
5	.12	50.7	.84	.82	D 16	articles
5	.11	49.7	1.10	1.20	k 14	sum spokEng
5	.10	60.1	.80	.79	E 9	academic
5	.03	27.8	1.01	1.09	J 19	main ideas
5	−.45	52.1	1.37	1.37	I 7	schemata
5	−.34	59.4	1.08	1.07	A 17	journals
5	−.30	60.3	1.04	1.03	B 15	editorials
5	−.28	56.5	.83	.86	i 20	idioms
5	−.27	46.5	1.15	1.11	d 24	scanning
5	−.25	47.9	1.02	1.04	j 21	quick read
5	−.25	58.7	.77	.77	G 2	predict
5	−.17	42.3	1.24	1.06	f 23	skimming
5	−.12	53.8	1.08	1.08	a 6	grammar
5	−.09	49.9	.94	.91	c 5	vocabulary
5	−.01	49.4	.96	.99	e 8	sentence

What we can conclude from this factor analysis is that reading ability means students read authentic materials, comprehend the materials, obtain the main idea and then draw some conclusions. A reading test measures how much students comprehend the material, and how much students' conclusions agree with the test makers' intention. In the process of comprehending the material, they will use skimming and scanning ability and the prompt context-grasping ability.

For testing reading ability, the material should be authentic and the end product can be measured by using M-C tests, Cloze tests, summary tests (in English or in Japanese) or translation tests. The measurement should be objective.

For teaching reading, we should take into consideration students' handling of the materials. The skimming and scanning and the prompt context-grasping skills (abilities) are both important to comprehend the materials quickly, without deviating from the main idea. Therefore, these two skills (abilities) should be strengthened in the classroom practice.

2. Comparison between Japanese and Non-Japanese teachers attitudes towards the items by mean scores.

Table 6

	M	SE	N
J	58.56	.89	43
NJ	54.71	.99	31

Table 6 shows that Japanese teachers use the higher point in the 1-4 scale than NJ teachers do.

Table 7

Category	Observed Count	
	J	NJ
1	42	120
2	275	185
3	481	284
4	363	248

Table 7 demonstrates that Japanese teachers tend to avoid the lowest point in the scale while NJ teachers use the scale widely. This is also supported by the mean score comparison mentioned above.

3. Conclusion

What we obtained from the Rasch analysis by focusing on the linearity idea (teachers' latent thought towards items and their actual opinion on the questionnaire) is that reading for testing has two components: contents and skills. Contents should be authentic materials because students use skills to comprehend or draw conclusions from the contents. The eventual scores given by the tests are the end product of their comprehension. Thus, we look at the process of reading by following three factors: first, students are given authentic materials; second, they struggle with the materials to comprehend, and try to use their skills; finally, their scores are calculated.

The way of testing reading ability (or the task which will be assigned to students) can vary depending on practicality or feasibility. However, in any case, ability should be measured objectively.

In addition, we were able to grasp the difference between Japanese

teachers' rating scale usage and NJ teachers' rating scale usage. Non-Japanese teachers are apt to use the scale more widely than Japanese teachers do.

Note: This research was supported in part by Tokyo Keizai University under Research Grant APC01-97.

Acknowledgements:
I am grateful to Dr. Benjamin D. Wright and Dr. John Michael Linacre for their valuable comments on this paper.

Reference

Bode, R.K. and B.D. Wright. (1999). Rasch measurement in higher education. *Higher Education: Handbook of Theory and Research*, Vol. XIV (pp. 287-316). New York: Agathon Press.

Nakamura, Y. (1998). Components of reading ability. *Educational Studies*, 40: 259-281. International Christian University: Tokyo.

Wright, B.D. (1997). A history of social science measurement. *Educational Measurement: Issues and Practice*, Winter 1997, 33-45 & 52.

Wright, B.D. (1999). Fundamental measurement for psychology. In S. Emretson, and S. Hershberger (Eds.). *The New Roles of Measurement: What Every Psychologist and Educator Should Know* (pp. 65-104). Hillsdale NJ: Lawrence Erlbaum Associates.

Wright, B.D., Masters, G.N. (1982). *Rating Scale Analysis: Rasch Measurement*. Chicago: MESA Press.

Wright, B.D. and Stone, M.H. (1979). *Best Test Design: Rasch Measurement.* Chicago: MESA Press.